Patent Application of

Rajeev Sharma,

Rahul Khare,

Mohammed Yeasin, and

Satish Mummareddy

for

# TITLE: BINARY AGE CLASSIFICATION OF HUMANS FROM DIGITAL IMAGES

## CROSS-REFERENCE TO RELATED APPLICATIONS

This application is based on and claims priority to U.S. Provisional Application No. 60/421,717, filed October 28, 2002, which is fully incorporated herein by reference.

## FEDERALLY SPONSORED RESEARCH Not Applicable

## SEQUENCE LISTING OR PROGRAM Not Applicable

## BACKGROUND OF THE INVENTION--FIELD OF THE INVENTION

This invention relates to the use of pattern recognition methodologies for developing a system for automatic categorization of a person from his image into a particular age category.

## BACKGROUND OF THE INVENTION

Age Classification has numerous applications and has the potential of not only enhancing the existing HCI system but can also serve as platform for passive surveillance (for e.g., alerting medical authorities if there is a accident in old age home). It can also be used for development of new HCI application (e.g., cigarette vending machines based on age verification), immersive computer games (for e.g., changing scenarios and multimedia content based on age category preferences), collecting retail business information (e.g., the number of children entering a retail store on a given day), image retrieval (for e.g., accessing all images belonging to babies), enhancing identity verification, and advertising (for e.g., focusing on a particular age group for selling a product).

To date there has been only one attempt to classify a person in an age category just from the facial information. U.S. Pat No. 5,781,650 to De Lobo describes an automatic feature detection and age classification method for human face in images. Their automatic age categorization system is based on finding a face in an image and locating the facial features such as nose, eyes, mouth, top of the head and chin. These features were then used to determine different T ratios (transverse ratios of distance

between eyes to the distance between line connecting the eyes from the nose/chin/top of head) that were then used for classification purposes. In the paper titled "Age Classification for Facial Images" by the same inventors Young H. Kwon and Niels De Vitoria Lobo, Computer Vision and Image Understanding, 74(1), pp. 1-21, 1991, they described their above patented method based on cranio-facial development theory and wrinkle analysis for age classification. In their invention, they did not use direct appearance information available from a face image to classify, instead they used geometric ratios obtained from the position of the facial features and presence of wrinkles.

Patent granted to Viola, US Pat No. (Application) US20020102024A1, describes a method for object detection using integral image representation of the input image. The object detector uses a cascade of homogenous classification functions or classifiers. Their invention defines a fast method for object detection using rectangular components defined by wavelets. The research paper titled "A Unified Learning Framework for Real Time Face Detection & Classification", Gregory Shakhnarovich, Paul Viola, and Baback Moghaddam, International Conference on Automatic Face and Gesture Recognition, 2002, performed gender and ethnicity classification using integral image. It calculates the integral image rather than classifying on basis of the face appearance. Furthermore, their system does not perform age classification.

U.S. Pat. No. 5,963,670 to P. R. Lipson et. al. , describes a method for classifying and detecting objects in images using a class model based on global

deformable templates. This method is based on building a class model in terms of a deformable template and cannot be applied for age classification.

In Andreas Lanitis, Chris J. Taylor and Timothy F. Cootes, "Towards Automatic Simulation of Aging Effects on Face Images", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 24, No. 4, April 2002, some work has been done in the field of simulating aging in facial images. In that paper the main aim was to make face recognition robust with respect to aging variations. Thus given the face of a person the face is "age normalized" before being used in face recognition. This method can be used for simulating ageing effects but does not address age classification.

In D. Micheal Burt, and David. I. Perrett, "Perception of age in adult Caucasian male faces: computer graphic manipulation of shape and colour information", Proceedings of the Royal Society of London, Vol.259, pp 137-143, 1995, a study was performed on the correlation between the perceived age and the chronological age. Though this study provides an insight into the ageing process, it does not deal with age classification of digital face images.

Patent granted to Player, US Pat No. (Application) US20020052881A1, shows an example of use of demographic information for customizing computer games and advertising. They did not show any method or system for extracting demographic information from images or videos.

SUMMARY

This invention deals with the age categorization of people from their low-resolution facial images into two groups.

The method according to this invention, comprise of a face detector module for identifying the faces within a digital image. The output of the face detector module is fed to a feature extraction module. The feature extraction module comprises of means for processing the face region to extract discriminating features by algebraic manipulation. These features are fed to classifiers to extract age category information.

The classifier module comprises of one or more classifiers that can work in serial or parallel or a combination to determine the age category for the person. Each classifier can make use of any known pattern recognition technique and be trained on a different set of features.

The two age categories can be obtained by any partition of the age spectrum into two groups.

The system consists of a set of visual sensors interfaced with a processing unit as the hardware. The digital images captured by the visual sensor are analyzed by face detector software installed on the processing unit. The output of this algorithm is fed to the classification system that decides the age category of the person.

DRAWINGS--FIGURES

FIG 1 – Age Classifier system overview

FIG 2 – Block diagram of the age classifier system

FIG 3 – Block diagram for training a classifier

FIG 4 – Block diagram for the parallel paradigm for classification

FIG 5 – Block diagram for the serial paradigm for classification

DETAILED DESCRIPTION OF THE INVENTION

An exemplary setup of the invention is now described in detail. The present invention detects the faces in the current scene and decides the age category of the persons. FIG 1 shows an exemplary hardware setup for the present invention. The hardware components of the system consist of an image capture device 101 and a processing unit 102.

In this exemplary embodiment shown in FIG. 1, a single firewire camera, such as the Pyro 1394 web cam by ADS technologies or iBOT FireWire Desktop Video Camera by OrangeMicro, or a USB camera, such as the QuickCam Pro 3000 by Logitech may be used as the image capture device 101. A computer system, such as the Dell Precision 420, with processors, such as the dual Pentium 864Mhz microprocessors, and with memory, such as the Samsung 512MB DRAM, may be used as the processing unit 102 in the exemplary embodiment. The processing software may be written in a high level programming language, such as C++ could be used in the exemplary

embodiment. FIG 2 shows the block diagram consisting of the main blocks used in the age category classification system.

An exemplary system is first fed the digital image 201 of the person as captured by the image capture device 101. The area of interest is cropped out from the scene by the face detector 202. The face detector algorithm as implemented in Henry A. Rowley, Shumeet Baluja, and Takeo Kanade, "Neural Network-Based Face Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, volume 20, number 1, pages 23-38, January 1998, could be used in the face detector 202 software.

The output of the face detector 202 software is a rectangular window containing the face of the person. Before the image is fed to the classifier, the image is passed through a pre-processing and feature extraction 203 process. Image processing algorithms such as histogram equalization and brightness gradient removal could be applied in order to present images of uniform brightness to the classifier. The brightness gradient removal algorithm consists of a method to remove any gradual changes of lighting across an image. Principal component analysis as explained in, Shree K. Nayar, Hiroshi Murase, Sameer A. Nene, " Parametric Appearance Representation", Chapter 6 in Early Visual Learning, Edited by Nayar and Poggio. Oxford University Press, 1996, and non-negative matrix factorization as explained in D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," Nature, vol. 401, pp. 788-791, 1999 etc, could be used. This representation of the image is finally fed to the binary age identification 204 system that decides the age category 205 of the person.

FIG 3 describes the steps in an exemplary case that could be followed for creating an accurate and generalized model classifier. In data collection facial images of people of different age groups were collected. All these images were appropriately labeled with the age category of the person in the image. These labels were used as ground truths to be used during the training 301 of the classifiers 302. This data set was divided into three parts – the training set, the bootstrapping set and the testing set all of them mutually disjoint. Histogram equalization and brightness gradient removal were used as part of the preprocessing for all the images in order to get rid of all brightness information that might affect the performance of the classifiers. Principal component analysis performed as a feature extraction. Different feature extraction methods can be used to generate different classifiers, for example the gray-scale value of the image pixels, and non negative matrix factorization. Besides increasing accuracy by removing very specific information about the images the feature extraction method also improves the computational speed of the classifier that is an important criteria for a real time classifier system.

The classifier 302 can be any pattern recognition algorithm. In this example support vector machines, C. Burges, "A tutorial on support vector machines for pattern recognition, data mining and knowledge discovery", vol. 2, pp. 121-167, 1998, can be used.

For the exemplary system, the collected data was divided into training data, bootstrapping data and test data. The classifier is trained 302 on the training data to create the primitive classifier 303. Then bootstrapping 304 is performed on the primitive

classifier 303. Bootstrapping 304 involves (i) testing the primitive classifier 303 on the bootstrapping data, (ii) separating out the misclassified data from the bootstrapping data, (iii) adding a fraction of the misclassified data to the training data, and (iv) training again to obtain the retrained classifier 305. The classifier performance 307 is determined by testing 306 the retrained classifier 305 on the test data. The above procedure is repeated for different values of parameters 301 to determine the retrained classifier for each set of parameter values. Classifier selection 308 is done by choosing the retrained classifier 305 with the best performance as the final classifier 309.

In order to improve the efficacy of the classifier either the parallel or the serial or a combination of the two paradigms could be used. The parallel paradigm as indicated in FIG 4 is based on the fact that examples misclassified by one classifier could be classified correctly by another thus giving a better overall accuracy if both the classifiers are used. The final classifiers 309 used in the bank of parallel classifiers 401 can vary either in the type of parameters used in the classifier or the type of feature extraction used for them. They could also differ in the type of classifiers used. Classifier decision fusion 402 is performed on the outputs of the final classifiers 309 in the bank of parallel classifiers 401 to obtain the age category 205. Another case of serial paradigm of classification could be one in which the first few stages of classifiers could be used as gross-level classifiers 501 followed by more fine-level classifiers 502 as indicated in FIG 5. Using the parallel and the serial paradigm simultaneously would give the best possible performance.

A particular exemplary case of age category classification could be a binary age category classifier using the serial paradigm for classification. In this example, the image from the camera is fed to the face detector software to detect the face in it. This face is then resized to the desired size and histogram equalization and brightness gradient removal is carried out on the image. Following the image processing the image is passed through a feature detector having a set of 100 basis vectors obtained from the training images thus giving a feature vector with 100 values. This is then fed to an age classifier. The final output of the age classifier gives the age category of the person as belonging either to the adult age category or the minor age category.